Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

**CSAS** | **SCCS**

Canadian Science Advisory Secretariat | Secrétariat canadien de consultation scientifique

Proceedings Series 2006/023 | Série des comptes rendus 2006/023

# Using Metadata Standards to Achieve Data Interoperability: Proceedings of a Technical Symposium and Workshop

# Utilisation des normes des métadonnées pour parvenir à l'interopérabilité des données : compte rendu d'un symposium et d'un atelier technique

Bedford Institute of Oceanography, Dartmouth, Nova Scotia

13-14 June 2006

R. Branton, T. Worcester, P. Lafond and D. Ricard

Institut océanographique de Bedford, Dartmouth (Nouvelle-Écosse)

Les 13 et 14 juin 2006

R. Branton, T. Worcester, P. Lafond et D. Ricard

Fisheries and Oceans Canada / Pêches et Océans Canada
Bedford Institute of Oceanography / Institut océanographique de Bedford
Dartmouth, Nova Scotia / Dartmouth (Nouvelle-Écosse)
B2Y 4A2 Canada

December 2006 | décembre 2006

Canada

## Foreword

This workshop was not carried out as a formal DFO Science Advisory Process; however, it is being documented in the CSAS Proceedings series as it presents some topics of interest related to the advisory process. The purpose of these proceedings is to archive the activities and discussions of the meeting, including research recommendations, uncertainties, and to provide a place to formally archive official minority opinions. As such, interpretations and opinions presented in this report may be factually incorrect or mis-leading, but are included to record as faithfully as possible what transpired at the meeting. No statements are to be taken as reflecting the consensus of the meeting unless they are clearly identified as such. Moreover, additional information and further review may result in a change of decision where tentative agreement had been reached.

## Avant-propos

Cet atelier n'a pas été tenu dans le cadre officiel du Processus de consultation scientifique de secteur des Sciences du MPO. Il est toutefois documenté dans la série des Comptes rendus du SCCS, car il couvre certains sujets en lien avec ce processus. Le présent compte rendu fait état des activités et des discussions qui ont eu lieu à la réunion, notamment en ce qui concerne les recommandations de recherche et les incertitudes; il sert aussi à consigner en bonne et due forme les opinions minoritaires officielles. Les interprétations et opinions qui y sont présentées peuvent être incorrectes sur le plan des faits ou trompeuses, mais elles sont intégrées au document pour que celui-ci reflète le plus fidèlement possible ce qui s'est dit à la réunion. Aucune déclaration ne doit être considérée comme une expression du consensus des participants, sauf s'il est clairement indiqué qu'elle l'est effectivement. En outre, des renseignements supplémentaires et un plus ample examen peuvent avoir pour effet de modifier une décision qui avait fait l'objet d'un accord préliminaire.

2007-01166/



# Using Metadata Standards to Achieve Data Interoperability: Proceedings of a Technical Symposium and Workshop

# Utilisation des normes des métadonnées pour parvenir à l'interopérabilité des données : compte rendu d'un symposium et d'un atelier technique

Bedford Institute of Oceanography, Dartmouth, Nova Scotia

13-14 June 2006

R. Branton, T. Worcester, P. Lafond and D. Ricard

Institut océanographique de Bedford, Dartmouth (Nouvelle-Écosse)

Les 13 et 14 juin 2006

R. Branton, T. Worcester, P. Lafond et D. Ricard

Fisheries and Oceans Canada / Pêches et Océans Canada
Bedford Institute of Oceanography / Institut océanographique de Bedford
Dartmouth, Nova Scotia / Dartmouth (Nouvelle-Écosse)
B2Y 4A2 Canada

December 2006

décembre 2006

© Her Majesty the Queen in Right of Canada, 2006
© Sa Majesté la Reine du Chef du Canada, 2006

ISSN 1701-1272 (Printed / Imprimé)

Published and available free from:
Une publication gratuite de :

Fisheries and Oceans Canada / Pêches et Océans Canada
Canadian Science Advisory Secretariat / Secrétariat canadien de consultation scientifique
200, rue Kent Street
Ottawa, Ontario
K1A 0E6

http://www.dfo-mpo.gc.ca/csas/

CSAS@DFO-MPO.GC.CA

Printed on recycled paper.
Imprimé sur papier recyclé.

Correct citation for this publication:
On doit citer cette publication comme suit :

DFO, 2006. Using Metadata Standards to Achieve Data Interoperability: Proceedings of a Technical Symposium, 13-14 June 2006. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2006/023.

## TABLE OF CONTENTS / TABLE DES MATIÈRES

**SUMMARY / SOMMAIRE** ..... **iv**

**INTRODUCTION** ..... **1**

**TECHNICAL SYMPOSIUM** ..... **2**

OPENING REMARKS AND KEYNOTE ADDRESS ..... 2

SESSION 1: RELEVANCE AND APPLICATION ..... 4

SESSION 2: DISCOVERY AND INTEROPERABILITY ..... 7

SESSION 3: AUTHORITY AND ACCESSIBILITY ..... 11

CONCLUDING REMARKS ..... 13

PARTICIPANT FEEDBACK ..... 14

ACKNOWLEDGEMENTS ..... 15

REFERENCES ..... 15

**NATIONAL DFO WORKSHOP** ..... **16**

SUMMARY OF DISCUSSION ..... 16

CONCLUSIONS ..... 21

**APPENDICES** ..... **22**

Appendix 1. Table of Abbreviations ..... 22

Appendix 2. List of Symposium Participants ..... 23

Appendix 3. List of DFO Workshop Participants ..... 26

Appendix 4. Symposium Agenda ..... 27

Appendix 5. Biographies of Presenters ..... 29

Appendix 6. Summary of Follow-up Survey Results. ..... 33

## SUMMARY

Integrated and ecosystem-based management rely heavily on information products (e.g., ecosystem overview and assessment reports) derived from the synthesis of a wide variety of data being collected and managed by others. Metadata (information about data) are beyond a doubt the most critically important factor in ensuring that data are used effectively beyond satisfying the initial investigations. To this end, a technical symposium and workshop were held to address the use of metadata standards to achieve data interoperability on June 13-14, 2006. Welcoming remarks were given by Dr. Peter Smith as acting director of the Bedford Institute of Oceanography and a keynote address was given by Dr. Fred Grassle, founding director of the Institute of Marine & Coastal Sciences at Rutgers State University. The symposium consisted of fourteen oral presentations organized into three sessions on metadata: 1) Relevance and Application 2) Discovery and Interoperability 3) Authority and Accessibility. Each session opened with introductory remarks by a senior DFO scientist and closed with a moderated discussion and online demonstrations. Over 100 researchers, data managers and contractors from government, universities and the private sector participated in this symposium, many of whom were from out of town. Following the symposium, a workshop of the National Science Data Management Committee was held. The main objective of this workshop was to build on work that had been done at a previous Service Oriented Architecture and Interoperability Workshop and to discuss and agree on what DFO Science needs to do to create, manage and publish metadata about its data holdings and publications, taking into account collaboration with external stakeholders. In general, the participants agreed that the workshop was productive, provided a level-playing field on metadata and metadata standards, and clear recommendations and action items on how to proceed with metadata for Science data holdings.



## SOMMAIRE

La gestion intégrée et écosystémique repose largement sur les éléments d'information (p. ex. aperçus de l'écosystème et rapports d'évaluation) issus de la synthèse d'un vaste éventail de données recueillies et gérées par d'autres. Les métadonnées (information sur les données) sont assurément le facteur le plus essentiel à l'utilisation optimale des données au-delà de la réponse aux interrogations initiales. Les 13 et 14 juin 2006, on a tenu un symposium et un atelier techniques sur l'utilisation des normes des métadonnées pour parvenir à l'interopérabilité des données. L'allocution d'ouverture a été prononcée par Fred Grassle (Ph.D.), directeur fondateur du Institute of Marine & Coastal Sciences de la Rutgers State University. Le symposium se composait de quatorze exposés oraux, organisés en trois séances sur les métadonnées portant sur les thèmes suivants : 1) pertinence et application 2) découverte et interopérabilité et 3) autorisation et accessibilité. Chaque séance débutait par des observations préliminaires d'un scientifique principal du MPO et se terminait par une discussion dirigée et des démonstrations en ligne. Plus d'une centaine de chercheurs, gestionnaires de données et entrepreneurs des milieux gouvernementaux et universitaires ainsi que du secteur privé, beaucoup d'entre eux venant de l'extérieur de la ville, ont participé à ce symposium. Celui-ci a été suivi d'un atelier du Comité national de gestion des données scientifiques, qui visait principalement à donner corps au travail déjà effectué lors d'un atelier précédent sur l'architecture axée sur les services et l'interopérabilité, et à s'entendre, après discussion, sur ce que la Direction des sciences du MPO a besoin de faire pour créer, gérer et publier des métadonnées sur ses fonds de données et publications, en collaboration avec les intervenants de l'extérieur. En général, les participants ont jugé l'atelier productif et utile à l'uniformisation des règles régissant les métadonnées et les normes s'y appliquant. Ils ont estimé qu'il avait abouti à des recommandations et des mesures de suivi claires quant à la production de métadonnées sur les fonds de données de la Direction des sciences.

## INTRODUCTION

Metadata is a key component of any data management framework. It enables the cataloguing of data holdings and the discovery of these holdings by people interested in using them for further analysis.

A number of initiatives have been undertaken by various groups within Fisheries and Oceans Canada (DFO, see Appendix 1 for a complete list of abbreviations) over the years to define what metadata should be recorded about specific types of data holdings and where this metadata should be stored and managed for discovery. A national workshop on Service-Oriented Architecture (SOA) and Interoperability for DFO Science held in March 2006 in Mont-Joli recommended the following actions regarding metadata and cataloguing (DFO 2006):

1. Identify which specifications to be used by DFO Science for metadata.
2. Implement a catalogue of the Science archives and determine if the catalogue should be implemented as a Web service.
3. Prioritize and create the metadata needed to complete the archive catalogue.
4. For document management look at links to Electronic Knowledge Management Environment (EKME) or create a library.

As a parallel activity to this workshop, DFO Science in the Maritime Region undertook the organization of a Metadata Symposium and Workshop held at Bedford Institute of Oceanography (BIO), Nova Scotia, on June 13-14, 2006.

On June 13th and the morning of June 14th, a technical symposium was held to give providers and users of data on marine organisms and physical processes an opportunity to improve their understanding of how present day metadata standards and associated systems might be used to enhance the quality of their work. The symposium was principally organized by members of the Centre for Marine Biodiversity Technical Committee, was funded by the Department of Fisheries and Oceans, Martimes Region Science Director, and was attended by people from DFO and external organizations that collaborate with DFO in a number of initiatives (see Appendix 2 for a complete list of participants). The primary focus was on oral presentations and panel discussions by technology specialists presently engaged in the creation, operation and use of various open access metadata systems and or portals.

The workshop, held on the afternoon of June 14th, focused on DFO Science and was attended by members of the National Science Data Management Committee and a few additional DFO representatives. The complete list of attendees is included in Appendix 3. The main objective of the workshop was to build on the work that had been done at the SOA and Interoperability Workshop and to discuss and agree on what DFO Science needs to do to create, manage and publish metadata about its data holdings and publications, taking into account collaboration with external stakeholders.

A summary of the presentations, discussion and feedback from the technical symposium is provided, followed by a summary of the discussion, conclusions and recommendations of the internal DFO workshop.

## TECHNICAL SYMPOSIUM

On June 13th and the morning of June 14th, a technical symposium was held at BIO to give providers and users of data on marine organisms and physical processes an opportunity to improve their understanding of how present day metadata standards and associated systems might be used to enhance the quality of their work. A complete list of symposium participants is provided in Appendix 2, the agenda of the meeting is provided in Appendix 4, and biographies of presenters and session chairs can be found in Appendix 5.

## OPENING REMARKS AND KEYNOTE ADDRESS

[Click on titles to view presentations online]

***Using Standards to Achieve Data interoperability***
Peter Smith and Bob Branton
Department of Fisheries and Oceans, BIO

There is a need in government and beyond for data interoperability, which will be facilitated through the use of metadata standards. For example, the identification of Ecologically and Biologically Significant Areas on the Scotian Shelf will require that all available information (including local ecological knowledge) on distribution and abundance of organisms, oceanographic data and information on habitat and bottom types be gathered into, or made accessible through, a single source. Questions remain on whether to use a common access approach or data warehouses. Interoperability enables consistent and reliable use of data from disparate and often remote systems. Standards-based systems may be difficult to implement, but they do provide built-in interoperability. Other benefits include: efficiency and cost effectiveness through improved geomatics coordination in policy making and governance. However, issues such as data rescue, use of dialects, incorporation of model "data" and differing ontologies and nomenclature still need to be addressed. Many examples exist, including the South eastern Universities Research Association Coastal Ocean Observing and Prediction Prototype Lab and the Open Integrated Ocean Observing System Interoperability Test bed (www.openioos.org). In conclusion, timely environmental prediction requires rapid near-real time integration of many different types of data and model results. Such a task is impossible without the benefit of seamless interoperability provided by the strict use of metadata standards.

***Ocean Observation, Census of Marine Life, Ocean Biogeographic Information System (OBIS) & Ocean Data Interoperability***
Fred Grassle
Institute of Marine & Coastal Sciences (IMCS), Rutgers State University of New Jersey

As founding director of IMCS and principal investigator behind the Ocean Biogeographic Information System (OBIS), Dr. Grassle provided an extensive first hand view of new and emerging instrument-based ocean observing systems and their relationship to OBIS and the Census of Marine Life (CoML). Global images of sea surface temperature and dynamic topography from traditional sources such as passive imaging and active radar satellites were contrasted with regional water mass classification and frontal boundaries images for New Jersey and Long Island coastline derived from new enabling technologies such as high frequency radar networks, autonomous sub sea gliders and cabled seafloor observatories. Remote interactivity as provided by the Coastal Ocean Observation Lab (COOL) Center for Advanced and Sustained Technologies (CAST) at IMCS was explained to be the critical and defining element of these new technologies, with the ultimate objective in this case being a

long-term shelf-wide context for high resolution nested process studies such as near shore recirculation. The International OBIS Internet (IOBIS) portal at Rutgers was then shown to provide a series of authoritative species lists and fact sheets for these same coastal areas. The ORION and NEPTUNE cabled observatory projects in the geologically active NE Pacific basin were also described in terms of linkages to a number of CoML field projects namely: Global Census of Marine Life on Seamounts (CenSeam), Biogeography of Chemosynthetic Ecosystems (ChEss), and Natural Geography in Shore Areas (NaGiSA), International Census of Marine Microbial Life (ICOMM), all of which are part of the OBIS network. OBIS's main mission was explained to be one of providing the means to effectively publish primary data on marine species from all over the world, which today consisted of 9.5 million records on 61,000 marine species from 114 sources*. Primary factors given for OBIS success include end user support for OBIS through provision of metadata, particularly in the form of usage citations and comments on the operation of OBIS website and portal; as well as willingness to promote the value of OBIS to governments and funding agencies; and encouraging publication of new data through OBIS. Fred concluded by thanking everyone for attending the symposium and noting that his long time colleague and collaborator on the OBIS project, Dr. Karen Stocks now at the San Diego Super Computer Centre would provide a more technical description of OBIS system as a whole later in the program.

* As of Sep 25, 2006, data served via the OBIS Portal had increased to over 10 million records on 71,000 marine species from 142 sources.

***Three Oceans of Biodiversity***
Bob Branton
Department of Fisheries and Oceans, BIO

In November 2003, the Centre for Marine Biodiversity (CMD) had hosted a workshop somewhat like this one, although much smaller and principally focused on DFO Maritimes. A workshop report is given at http://www.marinebiodiversity.ca/en/standards/standards.html. The object at that time was to encourage publishing of marine biodiversity data using international standards. It was at this time that CMB had just begun to experiment with development of a Canadian OBIS Portal. Since then, there has been much progress. The OBIS Canada Portal is fully operation, with the CMB homepage appearing first on the list when asking Google about 'marine biodiversity'. Underlying this CMB has assisted with creation of FGDC and GCMD metadata entries for 15 Canadian marine data collections resulting in almost 700,000 records on 9,000 species being delivered to OBIS. There is more data to come very soon, principally from DFO Maritime and Pacific Regions. Carrying on from this symposium, BIO and CMB will be hosting a full scale international conference dedicated to 'Ocean Biodiversity Informatics' in November 2007. This will be second international conference by this name, the first being held at Hamburg Germany in November 2004. We are calling the conference OBI '07 for short and expect it to principally focus on issues of data quality assurance and creation of information products for decision makers. The conference announcement is available on line at http://www.iobis.org/meetings/meet2007 .

## SESSION 1: RELEVANCE AND APPLICATION

Session Chair: Peter Lawton, DFO

*Researchers and resource managers generally expect that ecological research such as definition of sensitive areas and routine publishing of status reports would be greatly facilitated if all of the available information on distribution and abundance of organisms, oceanographic data and information on habitat and bottom types were to be gathered into, or made accessible through, a single source. In this session we have four end users of scientific data provide insights into their expectations.*

***Gulf of Maine Ecosystem Indicators Partnership***
Ray Konisky
Gulf of Maine Council

The Gulf of Maine Ecosystem Indicator Partnership (ESIP) is a regional ecosystem indicators and reporting program for the Gulf of Maine/Bay of Fundy. While many indicator and reporting efforts exist within and encompass the Gulf of Maine, a gulf-wide program is currently lacking. ESIP, a new program of the Gulf of Maine Council on the Marine Environment and its partners, is a science-based initiative to leverage existing monitoring datasets into a comprehensive reporting system for regional decision-makers. Datasets that are Canadian Geospatial Data Infrastructure (CGDI) compliant represent a best first-cut of indicator data sources. The program will initially develop 1-2 indicators from each of six focus-areas: aquatic habitats, climate change, coastal development, contaminants, fisheries and aquaculture, and nutrients. ESIP has a dedicated program manager and is driven by a Steering Committee representing US and Canada marine-environmental interests. Programmatically, ESIP will facilitate efforts by the six focus-area working groups to identify, populate, and report regional indicators. As a jump-start to ESIP, a transboundary project funded by GeoConnections has been started to target regional nutrient and contamination monitoring results.

***State of the Environment Reporting***
Ken Frank (presented by Brian Petrie)
Department of Fisheries and Oceans, BIO

A recent review of the eastern Scotian Shelf ecosystem is given as an example of the use of a number of data sources to examine the dynamics affecting the fisheries. Biological, chemical and physical variables in approximately 60 time series of up to 50 years duration were assembled from databases along with other observations that reflected fishing effort and industrial activity. The analysis revealed systematic patterns in temporal variability with fundamental changes in the biological state and negative correlations between adjacent groups ordered as large predators, small pelagics, zooplankton, phytoplankton and nutrients (trophic cascades). The project profited from the availability of a number of databases and their annotations, and illustrates the need and utility of collating individual datasets.

***The Gulf of Maine Ocean Data Partnership – Building a Region-Wide Information System to Support Gulf of Maine Resource Management***
Tom Shyka
Gulf of Maine Ocean Observing System

How do you find, access and integrate geospatial data related to the Gulf of Maine to use for resource management decision making? The Gulf of Maine Ocean Data Partnership (GoMODP) hopes to simplify the answer to that question in the near future.

The GoMODP is comprised of 21 organizations that collect and manage environmental data within the Gulf of Maine and its watershed. Members include federal, state, provincial, university and research organizations in the US and Canada. Most of the data collected by the partners has a geospatial component to it and could eventually be used within a GIS framework to support resource management activities. The goal of the partnership is to make each partner's long term datasets discoverable, accessible, and eventually interoperable through tools available on the internet. The partnership intends to use standards and protocols already in use by the various disciplines represented wherever possible. With regard to geospatial data, the GoMODP is promoting the use of the Open Geospatial Consortium services, which allow data providers and users to dynamically share and integrate geospatial data over the web.

To fulfill that goal, partners have filled out detailed surveys regarding their data. This information, which is available to the other partners, will assist in creating guidance on developing interoperability. As a first step, metadata training and assistance are being provided to partners to aid in establishing a common set of practices in the design and publishing of metadata and to make data discoverable through the American Geospatial One Stop, The Global Change Master Directory (GCMD) and/or the Canadian GeoConnections Discovery Portals on behalf of the data partnership and the individual organizations.

Developing data discoverability, accessibility, and interoperability without making the task overwhelming for the partners is challenging for such a diverse set of organizations. Each year a work plan is established with clear goals and a governing board and technical committee ensure that the goals are accomplished during the year. The partnership continues to add members and hold annual meetings.

***Metadata and the International Polar Year Data Policy: Why, What, and How it links to Sustainable Ocean Management and Beyond***
Falk Huettmann
University of Alaska

The International Polar Year (IPY) suggests a holistic approach to the science and management of Polar regions. It has a global and interdisciplinary outlook, and is meant to provide for progress and sustainability in society, and for leaving a global legacy. It inherently includes a digital Data Policy that banks on Open Access and Metadata.

Science-based, Adaptive Management of Natural Resources is widely perceived as being among the best professional techniques safeguarding a sustainable globe for future generations. Data can only be used efficiently for such exercises when they are freely available, e.g. provided online to the public, and when described with high-quality standards so that they are well known in their entirety. In the context of the IPY Data Policy, this presentation provided an overview of global standards, applications and International Organization for Standardization (ISO) formats for databases and their Metadata related to

Marine Environmental Impact Studies, Biodiversity, Wildlife and Habitats in the Arctic and elsewhere. The proposed IPY Data Portal, the Data and Information Service, and how they link with on-going global data initiatives such as OBIS and Global Biodiversity Information Facility (GBIF) were discussed. An outlook was given for how these standards and approaches are implemented by the IPY Data Policy, and how it will affect science and management for the global village and its citizens.

## *Discussion*

ISO abstract standards vs. profiles
The relative importance of ISO abstract standards and community profiles was discussed. Upon understanding that standards are very abstract and don't make sense to the average user unless applied to a specific community of practice, it was concluded the average end-user should focus on understanding community profiles.

DIF vs. FGDC
The merits and downsides of the two dominant metadata standards, the Directory Interchange Format (DIF) and the Federal Geographic Data Committee (FGDC) standards, were discussed. While comprehensive, the FGDC standard is intimidating and can be difficult to a new user. The DIF format is simpler but its simplicity limits the amount of information that it can store. To ensure long-term sustainability of the data collected, the International Polar Year (IPY) will use the FGDC standard.

Duplicates and overlap of data
Numerous questions about data duplications were raised. Participants wanted to know what happens if or when the same data appears more than once in data portals. The Gulf of Maine Ocean Data Partnership (GoMODP) has a metadata editorial board that ensures that no duplication of information occurs on their portal.

Importance of metadata to the user community
The importance of making sure that metadata standards are made clear to data providers was noted. Potential users of the available data will greatly benefit from having access to up-to-date standardised metadata. Data providers have a lot of knowledge about the data products that they have available and it can be easy for them to forget the important role of the metadata in making their products useable by the research community at large.

Technical compatibility does not equal biological compatibility
Achieving technical compatibility between data providers does not mean that biological compatibility is achieved. The level of detail required to pursue biological analyses might not be captured by metadata standards, especially with simpler standards such as DIF.

## SESSION 2: DISCOVERY AND INTEROPERABILITY

Session Chair: Bob Keeley, DFO

*Discovery portals like GeoConnections and the Global Change Master Directory: manage information provided by original data providers and allow end-users to search that information for data collections particular to their needs. Interoperability is the term used to describe the technical capacity to consistently and reliably use data from different and remote systems is and results from community-wide acceptance of data standards.*

### *The GeoConnections Discovery Portal*

Andrea Buffam
Natural Resources Canada

GeoConnections is taking a global approach to improving the discovery, access, use and sharing of geospatial information. Using the internet, they harness the power of geomatics by enabling sharing of interoperability spatially correlated data, where data is always obtained from the closest point to source. GeoConnections has evolved from a centralized warehouse to fully distributed networks, which are currently under development in 54 countries around the world. GeoConnections is mandated to build Canada's SDI through partnerships and provides the following services:

- Applications: builds applications that serve specifically targeted communities of practice,
- Content: provides the framework and other thematic data with attributes prioritized by these communities of practice,
- Technologies & Standards: maintains, operates and expands the core infrastructure and standards, as required by users, and
- Policies: supports consistent geomatics policy development federally and nationally to reduce duplication and improve the use of the CGDI.

Of interest to some may be the funding opportunities available through GeoConnections, what are related to metadata (including core technology and infrastructure development), as well as to directed innovation. The GeoConnections Discovery Portal was demonstrated, including the advanced search features, content editing, and metadata entry form.

### *Serving Fisheries and Ocean Metadata to Communities around the World*

Melanie Meaux
National Aeronautics and Space Administration (NASA)

NASA's Global Change Master Directory (GCMD) assists the oceanographic community in the discovery, access, and sharing of scientific data by serving on-line fisheries and ocean metadata to users around the globe. As of January 2006, the directory holds more than 16,300 Earth Science data descriptions and over 1,300 services descriptions. Of these, nearly 4,000 unique ocean-related metadata records are available to the public, with many having direct links to the data. In 2005, the GCMD averaged over 5 million hits a month, with nearly a half million unique hosts for the year.

Through the GCMD portal (http://gcmd.nasa.gov/), users can search vast and growing quantities of data and services using controlled keywords, free-text searches, or a combination of both. Users may now refine a search based on topic, location, instrument, platform, project, data center, spatial and temporal coverage, and data resolution for selected

datasets. The directory also offers data holders a means to advertise and search their data through customized portals, which are subset views of the directory.

The discovery metadata standard used is the Directory Interchange Format, adopted in 1988. This format has evolved to accommodate other national and international standards such as FGDC and ISO19115. Users can submit metadata through easy-to-use on-line and offline authoring tools. The directory, which also serves as the International Directory Network, has been providing its services and sharing its experience and knowledge of metadata at the international, national, regional, and local level for many years. Active partners include the Committee on Earth Observation Satellites, US federal agencies (such as NASA, the National Oceanic and Atmospheric Administration, and United States Geological Survey), international agencies (such as the Intergovernmental Oceanographic Commission, United Nations, and Japan Aerospace Exploration Agency) and organizations (such as ESIP, Integrated Ocean Observing System/Data Management and Communications, the Global Observing Systems Information Centre, Global Ocean Ecosystem Dynamics, OBIS, and GoMODP). In conjunction with OBIS, GCMD is implementing a new topic keyword called "BIOLOGICAL TAXONOMY" based on OBIS taxonomic categories thus providing additional taxonomy levels for improved data discovery of marine species data collections.

***COINAtlantic – Coastal and Ocean Information Network for the Atlantic***
Mike Butler
Atlantic Coastal Zone Information Steering Committee (ACZISC)

The ACZISC was established in 1992 to promote regional cooperation in Atlantic Canada with regards to coastal mapping, geomatics, and Integrated Coastal and Oceans Management. Membership is broad ranging, including the four Atlantic provinces, eight federal departments and agencies, community organizations, NGOs, the private sector and academia. The ACZISC has initiated the development of COINAtlantic to provide open access to regional data and information within the Canadian Geospatial Data Infrastructure, focusing on the needs of the Integrated Coastal and Oceans Management Community of Practice. Long-term goals include increased confidence of coastal stakeholders in coastal management decisions, and increased competitiveness of local marine geomatics enterprises in national and international markets. However, there are many challenges ahead. For example, COINAtlantic will have to determine the best method to deliver information to decision-makers and address the 'disconnect' between terrestrial and ocean data. Since challenges to data sharing are typically organizational in nature rather than technical, solutions will likely include changes in the attitudes and practices of users, providers and governments alike. The next steps for COINAtlantic are to secure commitment of a few key partners and data providers, and then to prepare a proposal to GeoConnections for the implementation of COINAtlantic.

***Towards Inter-jurisdictional Interoperability for a Sustainable Management of the St. Lawrence Ecosystem***
Joanne Hamel
Department of Fisheries and Oceans, Maurice-Lamontagne Institute

The evolution of information technologies and Web development tools has helped to improve access to various data sets, products and services for a wide range of users. Standards in many areas such as metadata, communication protocols and security play a key role in that respect. Catalogues, repositories and discovery portals also help users find the data they need. However, effort is required from data providers and end users to ensure that data are discoverable and easily accessible, understandable and useful. In addition, difficulties can be

compounded by a silo effect resulting from lacking integrated access to the various data sources and information systems.

In this context, the concept of Service-Oriented Architecture (SOA) was recently investigated by DFO from a systems interoperability perspective. A national pilot project was conducted which lead to the development of Web Data Services to facilitate the exploration of various heterogeneous data sets without requesting uniformity of formats, data models and technologies on the systems side, and without requiring specialized technical knowledge of the database structure and query mechanisms on the client side.

The same notions of standards, SOA and interoperability are core elements of other joint initiatives involving DFO, such as the St. Lawrence Global Observatory (SLGO), which brings together federal and provincial departments along with universities, and Research & Development and community organizations involved in the collection, management and diffusion of data about the St. Lawrence ecosystem. The SLGO vision is to provide efficient integrated access to the distributed data holdings of member organizations by fostering data, expertise and infrastructure sharing, networking and scientific partnerships. Such collaborative efforts involve well defined data management processes, roles, responsibilities, governance structure, service level agreements and the use of commonly agreed recognized standards.

***Mining New Information from Marine Data Archives: linking animal observations with time-synchronous oceanographic data***
Pat Halpin
Duke University

Significant effort has been expended in the development of observation programs for marine mammals and other critical management species. Much of this data is collected by government agencies and research laboratories in independent annual surveys with little attempt to synthesize or add value to these data archives. The OBIS-SeaMap program has collected more than 165 data sets on marine mammals, seabirds and sea turtles as a member of the larger OBIS and Census of Marine Life community. The group has developed a pilot program to merge archival observation data with time-synchronous oceanographic data to develop statistical habitat models that are directly useful for marine mammal avoidance questions. This processing allows for many years of previously collected data to be used in a new context and application. The presentation focused primarily on the workflow of processing time series data and the development of spatial and statistical techniques for extracting information from this composite data.

## *Discussion*

GeoConnections
The role of individual users in data entry versus data entry by designated data managers within GeoConnections was discussed. GeoConnections encourages the use of data managers that are responsible for numerous metadata entries; however, there may be some difficulty in keeping track of these people.

Lost in translation
Confusion over whether or not something is "lost in translation" when the contents of GeoConnections and GCMD are synchronized was clarified. Metadata flows one way, which is from the original provider through various conversions to the end-user. Anyone can report port problems, but only the originating provider is allowed to make changes. Although some information may be lost in translation, user discovering metadata at other than the originating

server can readily follow hyper links back to the original. Duplication of metadata entries may also be a problem. It is important to verify that a metadata entry does not already exist when submitting a new metadata record.

### *Demonstrations*

François Létourneau – Defence Research and Development Canada - Geolap
Mr. Létourneau showcased Geolap, a web application designed to facilitate the storage and recovery of geospatial data such as topographic maps. The software presented is currently in testing phase.

Mélanie Meaux – GCMD
Mélanie Meaux gave a demonstration of the Global Change Master Directory portal, including retrieval of underlying XML data using the GCMD Open Application Interface (API) facility.

Pat Halpin – OBIS-SeaMap and SERDP
Dr. Halpin presented two portals, OBIS-SeaMap and Strategic Environmental Research and Development Program (SERDP), that include map mapping services as well as "data products" that capture the workflow required to make products from raw data.

## SESSION 3: AUTHORITY AND ACCESSIBILITY

Session Chair: Ellen Kenchington, Centre for Marine Biodiversity

*Access portals like the Ocean Biogeographic Information System provide the actual means for getting data to the end-user. With public system such as OBIS, end-users are given access to standardized quality controlled results from a wide variety project database(s). Authority services like the Integrated Taxonomic Information System (ITIS) provide the means for names (species, locations, etc ...) to be unambiguously defined by providing them with the original source or authority for a given name.*

***The Ocean Biogeographic Information System: status, prospects, and challenges***
Karen Stocks
SanDiego Supercomputer Center

OBIS is an international federation of marine data providers working together to make their species distribution data available and interoperable through a singe web portal at www.iobis.org. At present, OBIS is serving over 9 million data points from over 100 datasets internationally, as well as connecting data to mapping and range prediction tools. As it grows, OBIS will be useful for such scientific issues as predicting climate change impacts on biodiversity, mapping hotspots of diversity in the oceans, and evaluating the level of sampling in different marine habitats and regions to guide future research.

***GeoPortal Web Map Server***
Don Vachon
Marine Environmental Data Services

Information was presented on the development of an International Standard based on ISO TC211, as well as the current S-57 levels of metadata and how they map to S100. The concept of Registries/Registers was introduced, and the way that GeoPortal is currently supporting metadata was described. The question, "what could it do more?" was addressed. Finally, a tour of the GeoPortal Web Map Services services/browser content/ functionalities was provided, using fisheries and oceanography examples.

***Biological Names, Metadata and Globally Unique Identifiers***
Guy Baillargeon
Agriculture & Agri-Food Canada

Taxonomic and vernacular names are key carriers of biological information. Names themselves are metadata for physical and digital objects (specimens, observations, surveys, images, sequences, publications). Taxonomic initiatives such as the Catalogue of Life – driven by Species 2000 and the Integrated Taxonomic Information System – are making substantial progress towards a synonymized checklist indexing the world's known species. However, names alone are not sufficient as identifiers in distributed databases as they lack stability and uniqueness. The Taxonomic Database Working Group and Global Biodiversity Information Facility are proposing to standardize on the use of Life Science Identifiers in setting up a system of globally unique identifiers to identify and access data objects on the Web, and metadata about them.

***The Atlantic Reference Center***
Lou Van Guelpin
Huntsman Marine Science Centre

The process leading to the sharing of biological data is based on specimens and standards. Biological data are obtained through the costly collection of specimens. The value of these expensive specimens will be maximized through the use of authoritative standards in each step of the data sharing process following collection: identification, taxonomy, archiving, digitizing, and transformation to formats for sharing via standardized metadata and portals. These standardized data can be used in-house or by others for research or to develop products such as regional species lists and species information systems.

The Atlantic Reference Centre (ARC) is a museum of Canadian Atlantic organisms, and a principal repository for specimens collected by DFO Maritimes. The ARC museum database is a source of biodiversity information to be shared, and is the basis of biodiversity information products. Therefore, the ARC is an institution that practices all stages of the data sharing process based on specimens and standards. Some of the pitfalls encountered and lessons learned involve taxonomic standards, data quality, data transformability, and the importance of accurate metadata.

***Creating a Marine Geospatial Data infrastructure at BIO***
Doug Gregory
Department of Fisheries and Oceans, BIO

Scientists, developers, data providers and informatics specialists at the Bedford Institute of Oceanography have been working towards the development of an infrastructure to support the delivery and dissemination of marine geospatial data. This presentation described some of these initiatives including provision of framework data, web-based data query applications, desktop user tools and the IT infrastructure to support the overall effort.

## *Discussion*

DFO participation in the Federal Biodiversity Information Partnership
The reasoning behind DFO's lack of involvement in the Federal Biodiversity Information Partnership was discussed. It was felt that the emphasis of DFO's contribution to the biodiversity community would be through OBIS.

Data below species-level
Whether existing data portals and the current data standards were suitable for sharing data below the species level was another topic of discussion. For example, what can be done to present data about different populations of a given species? Expansion of the OBIS schema to better capture observational data and provide information on populations and stocks was suggested.

Data for organisms that do not neatly fit the species concept
Methods to effectively capture and share data for living organisms that don't neatly fit the species concept was an issue of concern. For example, numerous bacteria and viruses are better described by "strains" than by "species". While the Catalogue of Life has a database for viruses, the database structure of ITIS is unlikely to be modified and will remain applicable only to names and taxonomic ranks covered by the botanical and zoological codes of nomenclature. This includes "below species-level" ranks such as subspecies and varieties, but not populations or strains.

Limits on number of records retrieved
Confusion over whether or not data portals limit the number of records than can be retrieved by users was clarified in part by noting that the OBIS portal currently limits queries to 10000 records.

Data with large storage requirements
The existence of numerous programs that collect vast amounts of data at high resolution was raised as a potential issue of concern. With the advent of mass storage for imagery and video files, the amount of information associated with a research initiative can become very large. Whether such data would be suitable for global portals was unclear given their large storage requirements.

Indicators of portal's popularity/usage
While the sheer number of collections/datasets/records available on a portal can be impressive, it was suggested that this should not be the only indicator of a portal's usability and success. Another indicator could be how many peer-reviewed scientific papers have been published using OBIS data.

The OBIS team is currently in the process of publishing a scientific paper about OBIS. This publication will then be used to cite OBIS and will facilitate the evaluation of OBIS-related publications.

## CONCLUDING REMARKS

Peter Smith thanked all participants for their input into the workshop.

Bob Branton thanked the participants and organisers of the workshop. He also announced the upcoming Ocean Biodiversity Informatics conference to be held at the Bedford Institute of Oceanography in November 2007.

The workshop was adjourned at 12:00.

## PARTICIPANT FEEDBACK

A follow-up survey was completed by 50 of the approximately 111 symposium participants. Graphic representation of these results is presented in Appendix 6. Of these respondents, roughly half were from DFO, 18% were from other government agencies (including Natural Resources Canada, Environment Canada, Agriculture and Agri-food Canada, Defence Research and Development Canada, and the Government of Nova Scotia), and 13% from universities. Almost half (45%) of respondents came from out of town; 41% attended from BIO. The majority of respondents identified themselves as researchers (48%) or data managers (26%).

Over 95% of respondents (47 of 49) felt that the symposium met their expectations, and participants had very positive things to say about the conference. In particular, participants appreciated the large turnout and attentive audience, the quality of the presentations and discussion, the venue and location (though perhaps a bit hard to find), diversity of topics and opportunity to network with the marine data management community.

Suggestions from individual participants by way of this questionnaire included: seating around tables so that notebooks and laptops can be used during (and after) the meetings, improved internet access, additional time for small group discussions, broader scope (i.e., more discussion of metadata standards in oceanography and not just biology), a more polished and complete program (received prior to the meeting), a title that was more reflective of the content (i.e., focus seemed to be on data portals and discovery rather than on metadata and increased interoperability once data are discovered), and an introductory discussion for those who were not yet familiar with the various standards and their uses (i.e., "standards 101").

In addition to asking participants for feedback on the symposium itself, the survey gathered additional information on metadata and other relevant service usage. The results of the survey indicated that all respondents use Google as a search engine on a regular basis; Yahoo, in comparison, tended to be used only occasionally. Most respondents use Geoconnections, with almost 75% claiming occasional to monthly use. GCMD and ITIS have been used by approximately 60% of respondents. 70% of respondents have never used Species 2000.

Eighty-eight percent of respondents indicated that they wished to receive notice and additional information on the upcoming OBI '07 workshop proposed for BIO next year.

## ACKNOWLEDGEMENTS

The organizing committee, Bob Branton, Victoria Clayton, Richard Eisner, Jennifer Hackett, and Tana Worcester, thank the presenters and participants for their time and effort. There were also a few others who were essential for the execution of the workshop and are also thanked: Dan Ricard, Jerry Black and Carolyn Harvey. Thank-you also to DFO Maritimes Science Branch for providing funding for this workshop and DFO's Oceans and Coastal Management Division for providing coffee and refreshments.

**(PRESENTERS)**
**Left to Right**: Lou Van Guelpen, Doug Gregory, Bob Branton, Falk Huettmann, Richard Eisner, Karen Stocks, Peter Lawton, Melanie Meaux, Andrea Buffam, Peter Smith, Guy Baillargeon, Victoria Clayton, Tom Shyka, Joanne Hamel.

## REFERENCES

DFO, 2006. National DFO Workshop on Service-Oriented Architecture (SOA) & Interoperability; March 28 – 30, 2006. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2006/024.

## NATIONAL DFO WORKSHOP

Following the technical symposium, a workshop of the National Science Data Management Committee was held at BIO on the afternoon of June 14th, 2006. The complete list of attendees is included in Appendix 3. The main objective of the workshop was to build on the work that had been done at a previous Service Oriented Architecture and Interoperability Workshop and to discuss and agree on what DFO Science needs to do to create, manage and publish metadata about its data holdings and publications, taking into account collaboration with external stakeholders.

## SUMMARY OF DISCUSSION

### *What are the units of metadata and which ones do we focus on initially (e.g., a collection, a dataset, or an observation/feature)?*

When looking at units of metadata, there is heterogeneity among DFO Science disciplines. Datasets and collections mean different things whether you are talking about hydrographic, biological, physical oceanography or environmental data. Datasets could be a national database, a series of surveys, a survey, a series of products or a specific product. In most cases, there are regional practices by discipline which define the unit of collection.

In hydrography, the structure has been established nationally across the Canadian Hydrographic Service (CHS) and follows the International Hydrographic Organization (IHO) specifications for products. CHS maintains metadata at the survey or source dataset level for its source data and at the feature, product and product collection level. IHO is currently developing a metadata profile based on the ISO 19115 standard, that will be part of the IHO S-100 specification, and CHS will adopt the specification once approved.

In physical oceanography, there is a long history of maintaining archives of data collected, often in databases. The unit of metadata is usually the database, qualified by the variables that are maintained in this database.

In marine biology, the most common unit of metadata is a series of surveys.

One of the current challenges is that it is often left to the project leader to determine what metadata should be recorded about a given project, as there are no clear guidelines regarding the unit of metadata as well as the level of details required. Furthermore, this metadata is often kept in an excel spreadsheet, electronic document or logbook and therefore not made available for general searches and discovery.

The following recommendations were agreed upon by the group:

1. Each region needs to identify the "big" archives/datasets that it holds and metadata will be recorded for each one. The definition of "big" will vary by discipline and region, but the purpose is to start with some key collections to have a broad range of metadata across all regions and disciplines.
2. Define national guidelines/best practices on metadata by science discipline.
3. Investigate how we can trace, through metadata, the relationship between value-added products and their source data, e.g. the climate database at BIO is based on the Marine Environmental Data Service (MEDS) archive. This could be through the use of parent/child records or other mechanisms.

***What attributes do we need and which specifications do we comply with?***

The discussion revolved mostly around the use of the FGDC – Content Standard for Digital Geospatial Metadata (CSDGM), the GCMD (a NASA initiative) and the ISO 19115 standard. The FGDC specification has been in used for many years for geo-referenced data and is currently the adopted specification for the Canadian Geospatial Data Infrastructure and for all U.S. government agencies. The ISO 19115 is an approved standard, but in order to be utilized, profiles need to be defined. Such a profile is currently being defined for North America and Natural Resources Canada (NRC) is a lead in this process. Profiles are also being defined for oceanographic and meteorological disciplines through work carried out by the WMO, but this is proceeding slowly. Following the workshop, Dr. Jean Brodeur from NRC, who is drafting the profile, confirmed that the North American profile will be finalized by the fall and he expects that it will be fully approved within a year. This profile will also become the new FGDC specification once approved.

The following recommendations were agreed upon by the group:

1. Finalize a DFO profile that would be based on the mandatory elements of FGDC but that would also include the mandatory elements from ISO. This high level profile would be the minimum that every discipline would need to comply with. Once the ISO 19115 North American profile is adopted, then DFO will move to the new profile.
2. Individual disciplines will define additional elements in their metadata specification, as required.
3. Educate DFO Science staff on Extensible Markup Language (XML) as it is the foundation to all metadata specifications.

***What classification/taxonomic system(s) do we use to qualify our metadata?***

Participants confirmed that different classification systems are required by discipline to assign keywords to metadata records. The GCMD keywords are the most widely used and have been adopted by NRC for the Discovery Portal. Furthermore, they have been translated to French. Therefore they should form the top level of the classification hierarchy. However, GCMD is prescriptive for the first three levels, leaving the fourth one to be user-definable. This is where each discipline needs to have a more detailed list of keywords. The following approach was proposed by discipline:

a. Marine Biology: there is a list of approximately 200 keywords that have been defined by OBIS for a high-level taxonomic classification (http://www.iobis.org/OBISWEB/ObisControllerServlet?searchCategory=/BrowseObisTaxCatServlet).
b. Physical oceanography: the MEDS data dictionary should be used (http://www.meds-sdmm.dfo-mpo.gc.ca/meds/About_MEDS/standards/login_e.asp or http://www.meds-sdmm.dfo-mpo.gc.ca/meds/About_MEDS/standards/login_f.asp ).
c. Hydrography: CHSDIR (the CHS internal catalogue) has a list of keywords.
d. Environmental sciences: does not appear to have a specific list of keywords defined.

A discussion ensued on how these keywords could be tested and it was proposed to pick a few disciplines and a few datasets and create metadata records to see if the keyword lists were appropriate.

The following recommendations were agreed upon by the group:

1. Use the GCMD list of keywords as the top levels of the classification system and supplement with more specific keywords by discipline as proposed above.
2. Conduct a pilot project with a few disciplines across disciplines to test the suitability of the keyword lists.
3. Ensure that the web page that will house search interfaces to the catalogues be fully qualified with appropriate keywords in Dublin core attributes, so that they are found easily through commercial search engines such as Google and Yahoo.

### *How do we create it?*

The discussion on metadata creation centered on two major topics: which tool(s) to use and language issues. A data entry tool is required as there will always be metadata records or components of records that need to be entered manually. However, it was suggested that automated creation of metadata records be used whenever possible, such as the project currently in progress supported by the National Science Data Managers Committee to automatically create metadata from oceanographic databases. With respect to the specific tool to be used, there are a few tools available, including the one currently deployed in the DFO GeoPortal that populates the catalogue, as well as tools such as docbuilder from NASA, which allows an end-user to create an XML file that can then be uploaded in a catalogue. docbuilder currently produces a DIF-compliant XML file, but there should be an FGDC-compliant version coming soon. (http://gcmd.nasa.gov/User/authoring.html)

It was agreed by all that metadata records that would be published externally needed to be produced in both official languages. This means that the tool used to create records must be available in both languages and also provide the user with keywords in the appropriate language. From a process perspective, it was proposed that records be created in one language and be validated before being translated to minimize the translation effort. The archive's custodian will be responsible to create the metadata record.

In summary the following recommendations were agreed upon by the group:

1. Create metadata automatically as much as possible from archives.
2. Different tools can be used to create the metadata record as long as they can output an XML file compliant with the DFO metadata specification based on FGDC.
3. Metadata will always be in both official languages.

### *How does it get validated?*

Validation is the process of ensuring that the documented metadata makes sense, that the descriptions are logical and easy to understand, that the right keywords have been applied and that all mandatory attributes have been provided. It was suggested that an editorial board be put in place to ensure that standards are followed and applied consistently. The issue of when to do translation to the second language came up and it was generally agreed that a metadata record should be developed in one language first, then validated and finally translated. Another level of validation will be required following the translation, to ensure that the translated text reflects accurately the meaning of the source record. It was also suggested to consult with the Communications Branch for their advice to develop a nationally consistent and stable process.

***Where do we store what?***

As described in section 2.1, various levels of metadata are captured by the different organizations within DFO Science. Some of this metadata, such as what is captured within the CHSDir database of CHS, is of an operational nature and is not intended for discovery and publication.

The NRC Discovery Portal is the primary metadata catalogue in the CGDI. Many DFO data holdings are currently maintained in the Discovery Portal, though some certainly need refreshing. There is also a synchronization facility between the Discovery Portal and the NASA GCMD catalogue to allow selected information to be duplicated in both catalogues for easier discovery. Some DFO data holdings maintained in the Discovery Portal are tagged with the OBIS keywords to allow them to be harvested by GCMD and made automatically available to the OBIS portal. This approach allows the metadata owner to maintain the records in only one place and make them available to many discovery sites.

DFO also currently maintains a national metadata catalogue through the GeoPortal. This catalogue, based on the M3Cat software from Intelec, has been used for the last three years to maintain metadata on the data available through the GeoPortal, on CHS data holdings as well as selected Science data holdings taken from SCIDAT. The catalogue is managed within an Oracle database. Its maintenance capability is web-based, therefore available from anywhere within DFO. An enhanced search interface was designed last year. The catalogue will be connected to the NRC Discovery Portal before the Fall to allow distributed searches and harvesting from the Discovery Portal.

The workshop participants agreed that there should be only one national discovery catalogue used by DFO and that the GeoPortal catalogue should be the one. This means that in order to satisfy the current requirement of OBIS, selected records will need to be maintained in GeoPortal and then harvested either by the Discovery Portal or by GCMD directly. It is also expected that metadata for the big archives should be maintained in the GeoPortal but pushed to the Discovery Portal for direct discovery through their search interface. Further review of the Discovery Portal harvesting and distributed query facilities will be required to ensure that the metadata maintenance task is minimized while the exposure of the data holdings is maximized. Additionally, as the new Open Geospatial Consortium Catalogue Service Specification gets implemented in various catalogues, the DFO Catalogue will need to provide an interface compliant with the new specification.

Currently, the DFO Catalogue supports the z39.50 protocol which is the most common protocol used, but which will change with the new catalogue specification.

In summary the following recommendations were agreed upon by the group:

1. The catalogue facility currently provided through the DFO GeoPortal will become the national metadata catalogue for DFO Science data holdings.
2. Metadata will be maintained in only one place but will be exposed to external catalogues, on a selective basis, using automated means provided by these external catalogues.

### *How do I make it available?*

The discussion on storage already covered the means to make the metadata available. However, the element of data sensitivity and security was brought up. Some metadata records will need to be flagged as sensitive, which means that they won't be accessible for discovery by the public. It was suggested that a component of the metadata record that would be mandatory is one that declares if the data can get public exposure or be restricted only to DFO exposure.

While the current GeoPortal catalogue facility is planned to be used, it was also recommended that the search interface not only be available through the GeoPortal but that there also be a branded search interface for DFO Science. Other than general look branding, the only difference in the search interface would be in the list of values for keywords, which would come from the agreed upon classification/taxonomic systems for Science data holdings.

There should also be ongoing discussions between DFO Science and other groups that maintain Science metadata catalogues to leverage what has been done by these other groups where possible and to share experiences and knowledge.

### *How do we manage it?*

It was agreed that the archive custodians own the metadata, and are therefore responsible for ensuring that it is always up to date. It is also critical to have an editorial board for quality control and develop a system of validation checks to identify records that may not be up-to-date, for example when the name of the contact person is no longer valid.

The flow could therefore be as follows:

1. Data archive custodian prepares metadata either directly in the DFO Catalogue (the current GeoPortal catalogue) or through another tool or automated process that produces an XML file and is then uploaded to the DFO catalogue. At that point the metadata record is in pending approval mode and is not available for discovery.
2. The Editorial Board reviews the metadata record and approves or sends comments back to the author.
3. Once the metadata record is approved, then the data archive custodian gets the record translated and submits the translated version to the Editorial Board who will then review for correctness (there may be potential savings brought by centralizing this translation service, as there will be a lot of common terms and similarities between records).
4. Once both the French and English versions of the metadata record are approved, then they are flagged as such in the catalogue and made available for discovery
5. If these records are required to be replicated to other catalogues, they will then be replicated via automated means.

On an annual basis, validation checks will be run by the DFO Catalogue administrator and a report will be produced. This report will be reviewed by the Editorial Board to determine updates required to specific records or general changes in procedures required.
More discussion will also be required on version control, to ensure that there is history maintained at the metadata level.

Additionally, as one of the early tasks in the implementation of a DFO Science metadata strategy it will be critical to find all existing DFO Science metadata records published in various metadata repositories (e.g. Discovery Portal, GCMD, GeoPortal, etc), review them and determine whether they are still valid or whether they should be deleted, updated or re-created.

## CONCLUSIONS

In summary, the participants agreed that the workshop was productive, provided a level-playing field on metadata and metadata standards, and clear recommendations and action items on how to proceed with metadata for Science data holdings. Some of these recommendations and action items are actually already identified in projects that were to be reviewed by the National Science Data Managers Committee immediately following the workshop. It is expected that most of the recommendations from the workshop can be implemented within this fiscal year.

## APPENDICES

### Appendix 1. Table of Abbreviations

ACZISC – Atlantic Coastal Zone Information Steering Committee
ARC – Atlantic Reference Centre
BIO – Bedford Institute of Oceanography
CGDI – Canadian Geospatial Data Infrastructure
CHS – Canadian Hydrographic Service
CMB – Centre for Marine Biodiversity
COIN – Coastal and Ocean Information Network
CSAS – Canadian Science Advisory Secretariat
CSDGM – Content Standard for Digital Geospatial Metadata
DFO – Department of Fisheries and Oceans
DIF – Directory Interchange Format
EKME – Electronic Knowledge Management Environment
ESIP – Ecosystem Indicator Partnership
FGDC – Federal Geographic Data Committee
GBIF – Global Biodiversity Information Facility
GCMD – Global Change Master Directory
GIS – Geographic Information System
GMBIS – Gulf of Maine Biogeographic Information System
GoMODP – Gulf of Maine Ocean Data Partnership
GoMOOS – Gulf of Maine Ocean Observing System
IHO – International Hydrographic Organization
IMCS – Institute of Marine and Coastal Sciences, Rutgers
IPY – International Polar Year
ISO – International Organization for Standardization
ITIS – Integrated Taxonomic Information System
MEDS – Marine Environmental Data Service
NASA – National Aeronautics and Space Administration
NGO – Non-Governmental Organisation
NRC – Natural Resources Canada
OBIS – Ocean Biogeographic Information System
OBIS-Seamap - OBIS-Spatial Ecological Analysis of Megavertebrate Populations
SCIDAT – Science Dataset Inventory
SDSC – San Diego Supercomputer Center
SDI – Spatial Data Infrastructure
SeaMap – Seabed Resource Mapping Program
SERDP – Strategic Environmental Research and Development Program
SLGO – St. Lawrence Global Observatory
SOA – Service-Oriented Architecture
XML - Extensible Markup Language

## Appendix 2. List of Symposium Participants

| Name | Affiliation | E-mail Addresses |
|---|---|---|
| | | |
| Adrienne Adamek | Aquatic Systems Groups, USM | |
| Guy Baillargeon | Agriculture and Agri-Food Canada | baillarg@agr.gc.ca |
| Lenore Bajona | DFO | bajonal@mar.dfo-mpo.gc.ca |
| Leslie Barton | DFO | bartonle@dfo-mpo.gc.ca |
| Pierre Belanger | DFO | belangerp@dfo-mpo.gc.ca |
| Nikki Bis-Saunders | Interpretation Resources | |
| Jerry Black | DFO | blackj@mar.dfo-mpo.gc.ca |
| Blair Dunn | DFO | dunnb@mar.dfo-mpo.gc.ca |
| Patrice Boivin | DFO | boivinp@mar.dfo-mpo.gc.ca |
| Shelley Bond | DFO | bonds@mar.dfo-mpo.gc.ca |
| Tony Bowdring | DFO | bowdringt@dfo-mpo.gc.ca |
| Bob Branton | DFO | brantonr@mar.dfo-mpo.gc.ca |
| Derek Broughton | Pointer Stop Consulting, Inc | broughtond@mar.dfo-mpo.gc.ca |
| Gary Bugden | DFO | bugdeng@mar.dfo-mpo.gc.ca |
| Mike Butler | ACZISC | michael.butler@dal.ca |
| Andrea Buffam | Natural Resources | BuffamAndreaNRCAN@dfo-mpo.gc.ca |
| Lesley Carter | Environment Canada | carterlesleyEC@dfo-mpo.gc.ca |
| Carla Caverhill | DFO | caverhillc@mar.dfo-mpo.gc.ca |
| Chiu Chou | DFO | chouc@mar.dfo-mpo.gc.ca |
| Pierre Clement | DFO | clementp@mar.dfo-mpo.gc.ca |
| Andrew Cooper | DFO | coopera@dfo-mpo.gc.ca |
| Lara Cooper | DFO | cooperl@dfo-mpo.gc.ca |
| Camille Coray | Dalhousie Biology Department | Camille.coray@dal.ca |
| Pat D'Entremont | Nicom | Pat.dentremont@nicomit.com |
| Rodolphe Devillers | Department of Geography, Memorial University of NL | rdeville@mun.ca |
| Len DiPenta | DFO | dipentam@mar.dfo-mpo.gc.ca |
| Penny Doherty | DFO | dohertyp@mar.dfo-mpo.gc.ca |
| Anna Fiander | DFO | fiandera@mar.dfo-mpo.gc.ca |
| Brad Firth | DFO | firthb@mar.dfo-mpo.gc.ca |
| Richard Eisner | DFO | eisnerr@mar.dfo-mpo.gc.ca |
| Paula Flynn | DFO | flynnp@dfo-mpo.gc.ca |
| Denis Gagnon | DFO | gagnondj@dfo-mpo.gc.ca |
| Jim Gale | DFO | galej@mar.dfo-mpo.gc.ca |

| Name | Affiliation | E-mail Addresses |
|---|---|---|
| | | |
| Steve Glavin | Nautical Data International, Inc. | |
| Fred Grassle | Rutgers University | grassle@marine.rutgers.edu |
| Michelle Greenlaw | Acadia Universit/Centre of Geographic Sciences | michelle.greenlaw@acadiau.ca |
| Doug Gregory | DFO | gregoryd@mar.dfo-mpo.gc.ca |
| Bob Groman | Woods Hole Oceanographic Institution | rgroman@whoi.edu |
| Jennifer Hackett | DFO | hackettj@mar.dfo-mpo.gc.ca |
| Sarah Hall | Environment Canada | hallsaragEC@dfo.mpo.gc.ca |
| Pat Halpin | Duke University | phalpin@duke.edu |
| Joanne Hamel | DFO | Hameljo@dfo-mpo.gc.ca |
| Carolyn Harvie | DFO | harviec@mar.dfo-mpo.gc.ca |
| John Holmes | DFO, MEAD | holmesj@pac.dfo-mpo.gc.ca |
| Tracy Horsman | DFO | horsmandt@dfo-mpo.gc.ca |
| Falk Huetttmann | EWHALE lab, IAB B&W, Univ. of Alaska-Fairbanks | fffh@uaf.edu |
| Sylvain Hurtubise | DFO | hurtubises@dfo-mpo.gc.ca |
| Diane Ingraham | Cape Breton University | diane_ingraham@capebretonu.ca |
| Anthony Isenor | DRDC Atlantic | |
| Todd Janes | DFO | janest@dfo-mpo.gc.ca |
| Anthony Joyce | DFO, Science Software System | joycea@dfo-mpo.gc.ca |
| Brian Keating | DFO | keatingb@dfo-mpo.gc.ca |
| Bob Keeley | DFO | keeleyb@dfo-mpo.gc.ca |
| Ellen Kenchington | DFO, CMB | kenchingtone@mar.dfo-mpo.gc.ca |
| Mary Kennedy | DFO | kennedym@mar.dfo-mpo.gc.ca |
| Ray Konisky | Gulf of Maine Council | rkonisky@comcast.net |
| Pierre Lafond | Holonics | plafond@holonics.ca |
| Peter Lawton | DFO | lawtonp@mar.dfo-mpo.gc.ca |
| Claudette LeBlanc | ACZISC | leblancc@ca.inter.net |
| Réjean Lebrun | Defence R&D Canada | Rejean.Lebrun@drdc-rddc.gc.ca |
| Francois Letourneau | Defence Scientist, Defence R&D Canada | Francois.Letourneau@drdc-rddc.gc.ca |
| Ed Light | Govt of Nova Scotia - GeoNOVA | |
| Peter Lockwood | Environment Canada | lockwoodpeterEC@dfo-mpo.gc.ca |
| Colin MacDonald | GeoNOVA Program Coordinator, GeoNOVA Secretariat | MACDONCW@gov.ns.ca |
| Kevin MacIsaac | DFO | macisaackg@mar.dfo-mpo.gc.ca |
| Denise McCullough | DFO | mcculloughd@mar.dfo-mpo.gc.ca |
| Melanie Meaux | NASA/GCMD | mmeaux@gcmd.nasa.gov |
| Eric Melanson | ESRI Canada | |

| Name | Affiliation | E-mail Addresses |
|---|---|---|
| | | |
| Michel Mitchell | DFO | mitchellm@mar.dfo-mpo.gc.ca |
| Dale Nicholson | DFO | nicholsond@mar.dfo-mpo.gc.ca |
| Robert Nowlan | DFO | nowlanr@dfo-mpo.gc.ca |
| Stephen Parsons | DFO, CHS | parsonssa@mar.dfo-mpo-gc.ca |
| Bernard Pelchat | DFO | pelchatb@dfo-mpo.gc.ca |
| Will Perrie | DFO | perriew@mar.dfo-mpo.gc.ca |
| Brian Petrie | DFO | petrieb@mar.dfo-mpo.gc.ca |
| Ted Potter | DFO | pottert@mar.dfo-mpo.gc.ca |
| Wayne Renaud | DND | |
| Daniel Ricard | Dalhousie University - FMAP | ricardd@mathstat.dal.ca |
| Dale Roddick | DFO | roddickd@mar.dfo-mpo.gc.ca |
| Sherrylynn Rowe | DFO | rowes@mar.dfo-mpo.gc.ca |
| Dave Senciall | DFO | senciall@dfo-mpo.gc.ca |
| Glyn Sharp | DFO | sharpg@mar.dfo-mpo.gc.ca |
| Tom Shyka | GoMOOS | tom@gomoos.org |
| Stephen Smith | DFO | smithsj@mar.dfo-mpo.gc.ca |
| Peter Smith | DFO | smithpc@mar.dfo-mpo.gc.ca |
| Tobias Spears | DFO | spearst@mar.dfo-mpo.gc.ca |
| Koren Spence | DFO | spencek@mar.dfo-mpo.gc.ca |
| Karen Stocks | San Diego Supercomputer Center | kstocks@sdsc.edu |
| Kohila Thana | DFO | thanak@mar.dfo-mpo.gc.ca |
| Scott Tomlinson | Canadian Federal International Polar Year Program Office | tomlinsons@dfo-mpo.gc.ca |
| Tony Turner | Natural Resources | TurnerTonyNRCAN@dfo-mpo.gc.ca |
| Bechara Toulany | DFO | ToulanyB@mar.dfo-mpo.gc.ca |
| Don Vachon | DFO, MEDS | Vachond@dfo-mpo.gc.ca |
| Lou Van Guelpen | Huntsman Marine Science Centre | arc@mar.dfo-mpo.gc.ca |
| Tony Walker | Jacques Whitford | |
| Louise White | GOM | |
| Brian Whitehouse | OEA Technologies Inc | bwhitehouse@eastlink.ca |
| Trudy Wilson | DFO | wilsont@mar.dfo-mpo.gc.ca |
| Nicholas Wolff | University Southern Maine - COML | nwolff@usm.maine.edu |
| Stewart Yee | Environment Canada | Stewart.Yee@ec.gc.ca |
| Kees Zwanenburg | DFO | zwanenburgk@mar.dfo-mpo.gc.ca |

## Appendix 3. List of DFO Workshop Participants

| Participant | Organization |
|---|---|
| LAFOND, Pierre | Holonics Inc. |
| SENCIAL, Dave | MPO/DFO-NFLD |
| SPEARS, Tobias | MPO/DFO-MAR |
| NICHOLSON, Dale | MPO/DFO-MAR |
| EISNER, Richard | MPO/DFO-MAR |
| O'NEIL, John | MPO/DFO-MAR |
| BRANTON, Bob | MPO/DFO-MAR |
| GREGORY, Doug | MPO/DFO-MAR |
| COSTELLO, Gerard | MPO/DFO-MAR |
| NOWLAN, Robert | MPO/DFO-GLF |
| HURTUBISE, Sylvain | MPO/DFO-QUE |
| HAMEL, Joanne | MPO/DFO-QUE |
| PELCHAT, Bernard | MPO/DFO-QUE |
| KEELEY, Robert | MPO/DFO-NCR |
| TRIVEDI, Tanya | MPO/DFC-C&A |
| HOLMES, John | MPO/DFO-PAC |
| BARTON, Lesley | MPO/DFO-PAC |

## Appendix 4. Symposium Agenda

# Using Metadata Standards to Achieve Data Interoperability

**Bedford Institute of Oceanography (BIO), Main Auditorium**

### Day One – June 13, 2006

**8:15** **Shuttle - leaves Ramada Inn for BIO**
**8:30** **Registration and Poster Set-Up**

### Opening Remarks

**9:00** **Peter Smith**, BIO – Opening Remarks
**9:15** **Fred Grassle**, Institute of Marine & Coastal Sciences, Rutgers – Keynote Address
**9:45** **Bob Branton**, BIO – Ocean Biological Information System (OBIS) Canada

**10:00** **Coffee Break** – hosted by Oceans and Coastal Management Division

### Session 1: Program Relevance and Application

**10:20** **Peter Lawton**, Saint Andrews Biological Station – Introduction to Session 1
**10:30** **Ray Konisky**, Gulf of Maine (GoM) Council – GoM Ecosystem Indicators Partnership
**10:45** **Brian Petrie** on behalf of **Ken Frank**, BIO – State of the Environment Reporting
**11:00** **Tom Shyka**, GoM Ocean Observing System – GoM Ocean Data Partnership
**11:15** **Falk Huettmann**, University of Alaska – Metadata and the IPY Data Policy
**11:30** **Moderated Discussion**

**12:00** **Lunch** – BIO Cafeteria

### Session 2: Discovery and Interoperability

**1:00** **Call to Order**
**1:05** **Bob Keeley**, Marine Environmental Data Services - Introduction to Session 2
**1:15** **Andrea Buffam**, Natural Resources Canada – Geoconnections Discovery Portal
**1:30** **Melanie Meaux**, NASA – Serving Fisheries & Ocean Metadata to the World
**1:45** **Mike Butler**, ACZISC – Coastal and Ocean Information Network for the Atlantic
**2:00** **Joanne Hamel**, Institut Maurice-Lamontagne – St. Lawrence Observatory
**2:15** **Pat Halpin**, Duke University – Mining New Information from Marine Data Archives
**2:30** **Moderated Discussion**

**3:00** **Coffee Break** – hosted by Oceans and Coastal Management Division

**3:30** **Hand-On Demonstrations and Poster Session**

**4:30** **End of Session – Closing Remarks**

**Day Two – June 14, 2006**

**8:15 Shuttle – leaves Ramada Inn for BIO**
**8:30 Late Registrations and Poster Set-Up**

**9:00 Opening Remarks and review of Day 1**

**Session 3: Data Authority and Accessibility**

**9:10 Ellen Kenchington**, Centre for Marine Biodiversity - Introduction to Session 3
**9:15 Karen Stocks**, SanDiego Supercomputer Center – OBIS
**9:30 Don Vachon**, Marine Environmental Data Services – GeoPortal Web Map Server
**9:45 Guy Baillargeon**, Agriculture & Agri-Food Canada – Biological names, metadata & globally unique identifiers.

**10:00 Coffee Break** – hosted by Oceans and Coastal Management Division

**10:20 Lou Van Guelpen**, Huntsman Marine Science Centre – Atlantic Reference Centre;
**10:35 Doug Gregory**, BIO – Creating a Marine Geospatial Data Infrastructure at BIO.
**10:50 Moderated Discussion**
**11:10 Hands-On Demonstration and Poster Session**

**11:45 Closing Remarks**

**12:00 End of Symposium**

## Appendix 5. Biographies of Presenters

### Guy Baillargeon

Guy Baillargeon is leading a small research group dedicated to the development of Web applications in the area of systematics and biodiversity at Agriculture and Agri-Food Canada. On behalf of the Federal Biodiversity Information Partnership, he is responsible for the technical development of the Canadian node (www.cbif.gc.ca) of GBIF. Guy is also responsible for the development of international versions of the Integrated Taxonomic Information System, shared by United-States, Canada and Mexico (itis.gbif.net). Since 2003, Guy is member of the Species 2000 Project Team, providing GBIF with the Catalogue of Life (www.species2000.org) checklist. In 2002-2004, he chaired the GBIF Participant Node Managers Committee (NODES).

### Bob Branton

Bob Branton had until recently been an assessment biologist in the Population Ecology Division at BIO with an interest in ocean biodiversity data and information which included participation in a number of internet initiatives including Gulf of Maine Biogeographic Information System (GMBIS), Ocean Biogeographic Information System (OBIS) and Gulf of Maine Ocean Data Partnership (GoMODP). At the beginning of 2006, Bob was given a special assignment to concentrate exclusively on management of ocean biodiversity data and information. Bob currently chairs the Centre for Marine Biodiversity Technical Committee (a.k.a. OBIS Canada), the OBIS Managers Committee and is vice chair of the GoMODP executive committee. Bob is a principal organizer of this symposium.

### Andrea Buffam

Andrea Buffam is the Senior Metadata Coordinator with the GeoConnections Discovery Portal. She has been working in Natural Resources Canada since 1991. Currently Andrea is with the Data Management and Data Dissemination Branch in the Earth Sciences Sector, working with GeoConnections to provide the data discovery and access component of the Canadian Geospatial Data Infrastructure. Andrea enjoys working collaboratively with national geospatial data specialists seeking greater interoperability and quicker access to data. She pursues opportunities to provide input into the direction of developing standards and the geospatial software technology as they evolve within the industry.

### Mike Butler

Mike is a Biological Oceanographer by training and currently Director of the Secretariat for the Atlantic Coastal Zone Information Steering Committee, located at Dalhousie University, Halifax, Nova Scotia and Director of the International Ocean Institute - Canada, also located at Dalhousie University

### Ken Frank

Ken Frank received his PhD from the University of Toledo and did his post-doc as a research associate at McGill University. He started work with DFO in the BIO Marine Ecology Laboratory in 1983. Since this time he has worked for the Marine Fish Division (1987-2002) and the Ocean Science Division (2002-present) at BIO. In addition to his work with DFO, he has also been an adjunct professor at Dalhousie University since 1989 and the Associate Editor of the Canadian Journal Fisheries and Aquatic Sciences since 1997. His research interests include: fisheries ecology, resource conservation, biogeographic theory, fisheries oceanography and marine ecosystem assessment.

### Fred Grassle

Fred Grassle is Professor of Marine and Coastal Sciences and Director of the Institute of Marine and Coastal Sciences (IMCS) at Rutgers University, The State University of New Jersey. Dr. Grassle founded IMCS in 1989 to consolidate University programs in marine and coastal sciences and develop an internationally known program of education and research in marine and coastal sciences. In addition to being Director of IMCS, Dr. Grassle is also the Director of the Fisheries Information and Development Center and New Jersey Chair of the NY and NJ Clean Ocean and Shore Trust.

### Doug Gregory

Doug Gregory has been actively involved in data management at BIO for over 30 years and is currently responsible for managing the Ocean Data and Information Services program for the Ocean Sciences Division. For the last decade, a major interest has been in providing open access to oceanographic products over the Internet.

### Lou Van Guelpen

Lou Van Guelpen is currently the Curator of Fishes and Collections Manager at the ARC and has been the ichthyoplankton taxonomist at the Huntsman Marine Science Centre/ARC for over 28 years. His primary research or administrative involvements include: taxonomy, life history, biogeography, laboratory processing procedures, and preservation of all life stages of Canadian Atlantic fishes; climate change impacts on fish, invertebrate, and plant populations; and internet integration of biodiversity information.

### Pat Halpin

Pat Halpin is currently the Gabel Associate Professor of the Practice of Marine Geospatial Analysis and Director of the Geospatial Analysis Program at the Nicholas School of the Environment and Earth Sciences, Duke University. Halpin directs the Marine Geospatial Ecology Lab research group at Duke University where he and his lab group are actively developing spatial analysis methods and technologies solving ecological and management problems in the marine environment. Halpin is a Co-principal investigator on the OBIS-SeaMap program (http://seamap.env.duke.edu ); a global observation database on seabirds, seaturtles and marine mammals. He serves on the OBIS International Steering Committee. Halpin is also a Co-Principal Investigator on the SERDP marine mammal habitat modeling program, develping new spatial-temporal modeling tools for marine species management applications. Halpin is currently involved in the development of new geospatial analysis tools to support emerging marine ecosystem-based management needs.

### Joanne Hamel

Joanne Hamel joined DFO in 1982 as a fisheries research assistant, first in Quebec and then at BIO. She became Assistant Coordinator and Webmaster for the Canadian Science Advisory Secretariat (CSAS) at DFO Headquarters in Ottawa in 1998 and has been involved in coordinating national projects and in developing the CSAS Internet site to make scientific advice and resource assessment information widely available. She joined the Maurice Lamontagne Institute in 2000 and is currently Head of the scientific data diffusion section of the Scientific Advice, Information and Support Branch. Her duties include being in charge of the St. Lawrence Observatory (http://www.osl.gc.ca) and managing Web development projects and multidisciplinary teams. Her areas of interest include the integration and accessibility of scientific information through the use of Web technologies and the adaptation of scientific information to the needs of various client groups.

**Falk Huettmann**

Falk received his M.Sc. in Munich and his PhD at the University of New Brunswick with the Atlantic Cooperative Wildlife Ecology Research Network (A.W. Diamond). He did a Killam PostDoc at the University of Calgary. Falk is now an Assistant Professor at the University of Alaska-Fairbanks. He works in his EWHALE lab with co-supervised students on GIS, databases, statistical biodiversity/habitat modeling and metadata on various topics and ecosystems world-wide. Falk is a frequent contributor of biodiversity and habitat data to various projects and programs for free public download to the global community.

**Robert Keeley**

Bob Keeley graduated from Dalhousie University with a MSc in physical oceanography. He has worked for DFO for almost 30 years, mostly at the ocean archive centre. He is a member of a number of national and international committees whose work centres around management of ocean data. Presently he is the chair of the Data Management Programme Area of the Joint Commission on Oceanography and Marine Meteorology, and activity sponsored by the Intergovernmental Oceanographic Commission and the World Meteorological Organization.

**Ellen Kenchington**

Dr. Ellen Kenchington is a Benthic Ecologist/Molecular Geneticist at the Bedford Institute of Oceanography with Fisheries and Oceans, Canada. Dr. Kenchington obtained her degrees from Dalhousie University (B.Sc., M.Sc.) and the University of Tasmania (Ph.D.). Current projects include investigations into sex determination in mussels, genetic structure of lobster and the impacts of fishing on the benthos. Ellen is Director of the Centre for Marine Biodiversity.

**Ray Konisky**

Ray Konisky has a Ph.D. in Environmental and Earth Sciences from the University of New Hampshire. He works on ecosystem indicator and modeling programs as an independent contractor for the Gulf of Maine Council on the Marine Environment, and as a staff scientist at the Wells National Estuarine Research Reserve in Wells, Maine. Ray recently completed a report entitled "A Regional Assessment of Salt Marsh Monitoring and Restoration in the Gulf of Maine".

**Peter Lawton**

Peter Lawton has been conducting field ecological research in marine benthic habitats of the northern Gulf of Maine since the late 1980's. Initially this work was conducted in support of invertebrate fisheries assessments and research on habitat requirements and sensitivities of marine invertebrates in relation to fishery production and interaction with other coastal development. Dr. Lawton's team has made extensive use of remote video techniques and *in situ* approaches such as SCUBA in a variety of project contexts. More recently, Peter Lawton with his colleague Mike Strong have been adapting these field survey approaches to support new demands for coastal habitat characterisation and marine biodiversity assessment in support of oceans management.

**Melanie Meaux**

Melanie Meaux is the Ocean and Antarctic Sciences Coordinator for NASA's Global Change Master Directory, located in Greenbelt, Maryland. She joined the group in September 2004. Prior to working with GCMD, she was a graduate student at the University of North Carolina at Chapel Hill where she received a Master of Science in Oceanography and was awarded the NASA Earth System Science Fellowship from 2001 to 2004.

### Tom Shyka

Tom Shyka is the Director of Program Development at the Gulf of Maine Ocean Observing System (GoMOOS). He received a BA in Biology and Environmental Science from Colby College in Maine and a MS in Marine Ecology from the University of Maryland. He has conducted research at various marine laboratories around the US and in the Caribbean. In addition, has worked for NOAA's National Marine Sanctuary Program and as an environmental consultant. In his current position at GoMOOS, Tom works with the various GoMOOS users (fishermen, commercial and recreational mariners, scientists, resource managers, and teachers) to help design useful information products that are available on the GoMOOS website. Additionally, he helps coordinates activities for the Gulf of Maine Ocean Data Partnership and is the program coordinator for the development of the Northeast Regional Association for Ocean Observing.

### Peter Smith

Peter C. Smith was trained in Physical Oceanography at the MIT-Woods Hole Oceanographic Institution Joint Program, where he received his PhD in 1973. Upon graduation, he came to BIO on a postdoctoral fellowship to pursue deep boundary current research, and two years later joined the research staff of the Coastal Oceanography Section. For the next two decades, he participated in and co-led various multidisciplinary research programs, including the Fisheries Ecology Program (1979-'83), GLOBEC Georges Bank Program (1990-'98), and the Canadian Atlantic Storms Programs I and II (1985-' 93). More recently, Dr. Smith has taken on science management roles, both within DFO and outside (Leader, Offshore Environmental Factors POL, Program on Energy Research and Development, 2000-present). His present scientific interests include surface waves and currents observation and modelling, ecosystem monitoring, and operational oceanography.

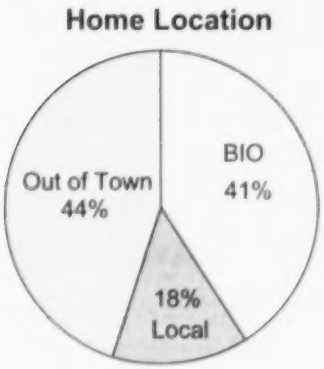
### Karen Stocks

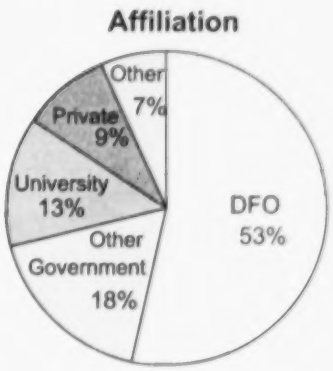
Karen Stocks was trained as a biological oceanographer, finishing her PhD research on estuarine invertebrate ecology at Rutgers University in 2000. She is now a researcher at the San Diego Supercomputer Center (SDSC) where she works at the interface of information technology development and ecological research. Her role at SDSC is to collaborate with computer scientists and programmers to help guide the development of new database and data integration technologies for use by marine biodiversity research and management. She developed the SeamountsOnline database (seamounts.sdsc.edu) and uses this resource for her research into large-scale patterns of biogeography on seamounts. She also participates in the development of OBIS.

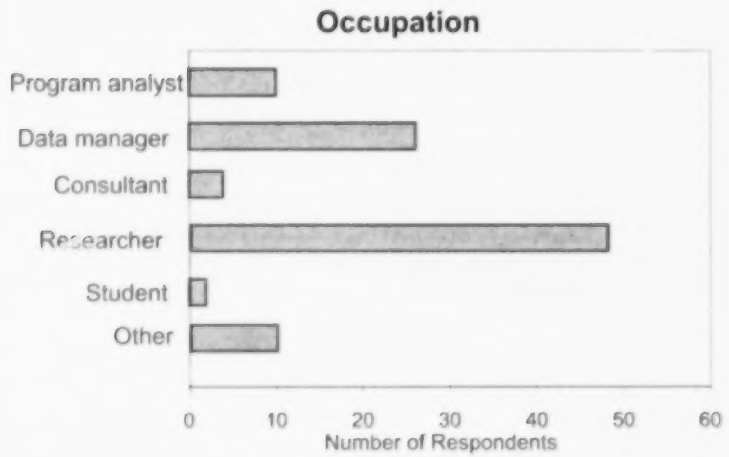
### Don Vachon

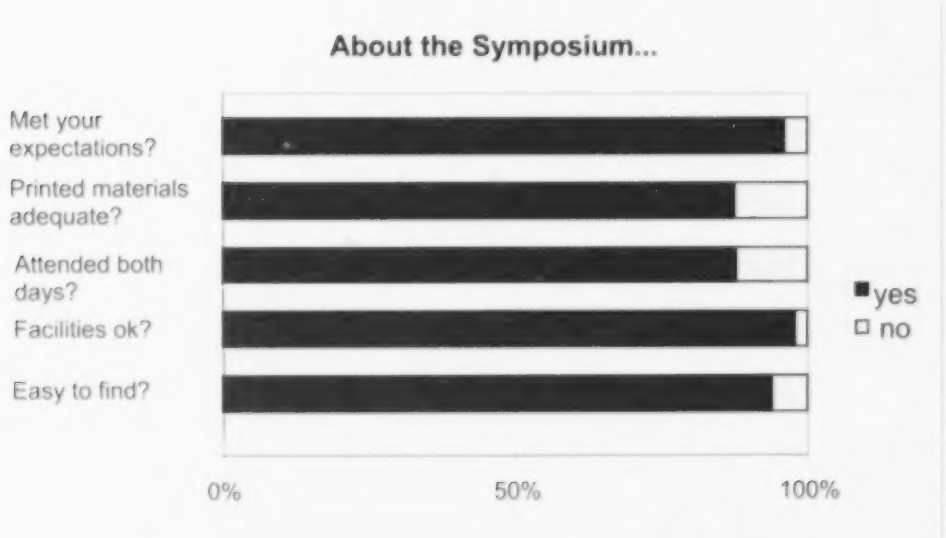
Don Vachon as Chief, Engineering Development and Geomatics Services, Canadian Hydrographic Service, has been involved with the development of GIS and geospatial databases for over 28 years and is currently implementing Open Geospatial Consortium-based services in his department through the GeoPortal Project. He is a member of the Working Group on S57 extensions for Edition 4.0 (S100) which integrates several of the ISO TC211 components including those relating to metadata. He is currently assigned to the new portfolio of Director Integrated Science Data Management.

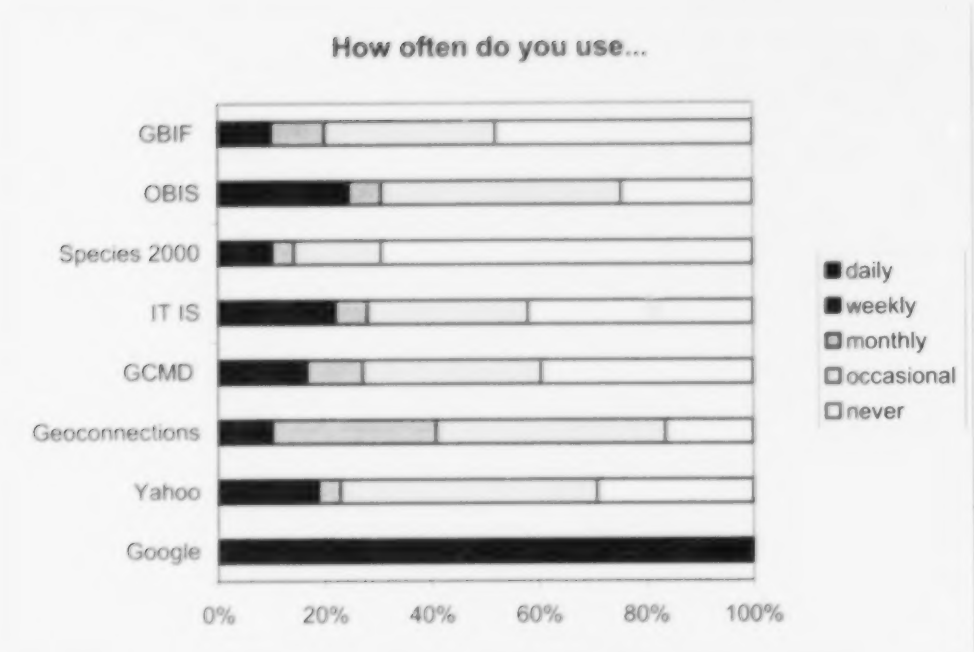
## Appendix 6. Summary of Follow-up Survey Results.

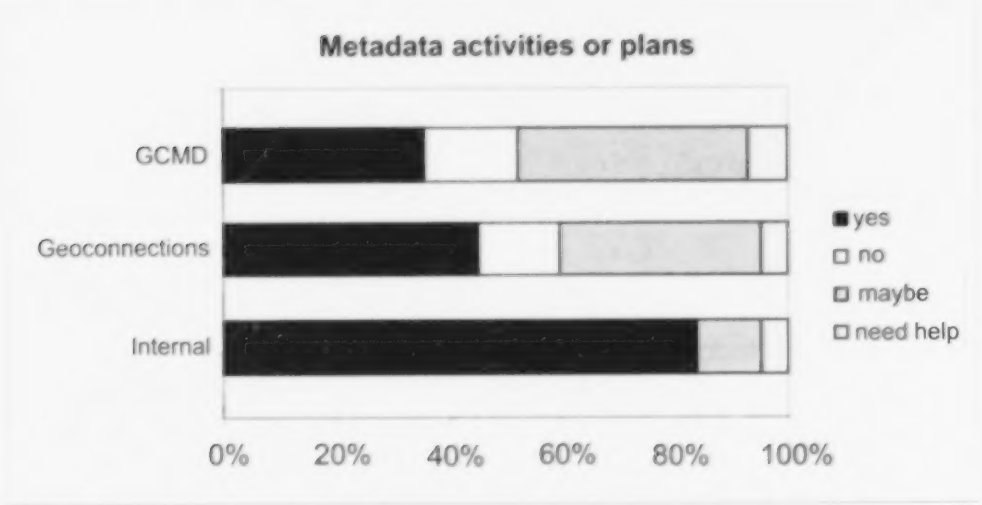



How often do you use...
GBIF
OBIS
Species 2000
IT IS
GCMD
Geoconnections
Yahoo
Google
0%
20%
40%
60%
80%
100%
daily
weekly
monthly
occasional
never


Metadata activities or plans
GCMD
Geoconnections
Internal
0%
20%
40%
60%
80%
100%
yes
no
maybe
need help